

```

1 CGAGTGCCT GCTGAGTCCT GTAGATAAAAG CCGCCAACCC CGGGGACTGG
51 TGTCTCTGA GTGACCGTGC AGCCACTTC CAGATCCTTC GGGCATTGG GAAGGGCAGC
101 CAGTGAACCT CGACCACTTC CAGATCCTTC GGGCATTGG GAAGGGCAGC
151 TTTGCAAGG TGTCATTGT GCAGAAGCGG GACACGGAGA AGATGTACGC
201 CATGAAGTAC ATGAACAAGC AGCAGTGCAT CGAGCGGAC GAGGTCCGCA
251 ACGTCTCCG GGAGCTGGAG ATCCTGCAGG AGATCGAGCA CGTCTCTG
301 GTGAACCTCT GTTACTCCTT CCAGGACGGAG GAGGACATGT TCATGGTCGT
351 GGACCTGCTA CTGGCGGGG ACCTGCGCTA CCACCTGCAG CAGAACGTGC
401 AGTTCTCCGA GGACACGGTG AGGCTGTACA TCTGCGAGAT GGCACACTGGCT
451 CTGGACTACC TGCGCGGCCA GCACATCATC CACAGAGATG TCAAGCCTGA
501 CAACATCTC CTGGATGAGA GAGGACATGC ACACCTGACC GACTTCAACA
551 TTGCCCCCAT CATCAAGGAC GGGGAGCGGG CGACGGCATT AGCAGGGCACC
601 AAGCCGTACA TGGCTCCGA GATCTTCCAC TCTTTGTCA ACGGGGGAC
651 CGGCTACTCC TTCGAGGTGG ACTGGTGGTC GGTGGGGGTG ATGGCTATG
701 AGCTGTCGCG AGGATGGAGG CCCTATGACA TCCACTCCAG CAACGCCGTG
751 GAGTCCCTGG TGCACTGTT CAGCACCGT AGCGTCCAGT ATGTCCCCAC
801 GTGGTCCAAG GAGATGGTGG CCTTGCTGCG GAAGCTCTC ACTGTGAACC
851 CCGGACACCG GCTCTCCAGC CTCCAGGACG TGCAAGGACG CCCGGCGCTG
901 GCCGGCGTGC TGCGGGACCA CCTGAGCGAG AAGAGGGTGG AGCCGGGCTT
951 CGTCCCCAAC AAAGGGCGTC TGCACTGCGA CCCCACCTTT GAGCTGGAGG
1001 AGATGATCCT GGAGTCCAGG CCCCCTGCACA AGAAGAAGAA GCGCTGGCC
1051 AAGAACAGT CCCGGGACAA CAGCAGGGAC AGCTCCCACT CGAGAATGA
1101 CTATCTCAA GACTGCTCG ATGCCATCCA GCAAGACTTC GTGATTTTA
1151 ACAGAGAAAAA GCTGAAGAGG AGGCCAGGAC TCCCGAGGGA GCCTCTCCCC
1201 GCCCCTGAGT CCAGGGATGC TGCGGAGCCT GTGGAGGAGC AGGCGGAACG
1251 CTCCGCCCCG CCCATGTGCG GCCCCATTG CCCCCTCGGCC GGGAGCGGCT
1301 AGGCCGGGAC GCCCGTGGTC CTCACCCCTT GAGCTGCTTT GGAGACTCGG
1351 CTGCCAGGAG GAGGCCATG GGCGGAGGCC TGGCATTACAC GTTCCCACCC
1401 AGCCTGGCTG GCGGTGCCA CAGTGGCCCG GACACATTG ACACCTCAGG
1451 CTCGTGGTGG TGCAAGGGAC AAGAGGCTGT GGGTGCAGGG GACACCTGTG
1501 GAGGGCATTG CCCGTGGGC CCCCAGGACCC GCCTAGATGG AGGAAGCGCT
1551 GCTGGCGCC CTCTTACCGC TCACGGGGAG CTGGGGCCAT GGATGGGACA
1601 GGAGTCTTG TCCCTGCTCA GCCCCGGAGGC TGTGCACCGC CCTCGTCACA
1651 AGGTGACCT TGCAAGCACAG GCGCGGGTG CCCCAGGCTC GGCTCAGTC
1701 TTGGAGGTCA AGGGCATGGG TTGGGGTAGT GGGTGGGGAG GTGAATGTTT
1751 TCTAGAGATT CAAACTGCTC CAGCAATTTC TGTATAGTTT TCACCTCTGA
1801 GAATTACAAT GTGAGAACCG CACAAAAAAA AAAAAAAA AAAAAAAA
1851 AAAAAAAAAA AAAA
(SEQ ID NO: 1)

```

FEATURES:

5'UTR: 1 - 192
 Start Codon: 193
 Stop Codon: 1300
 3'UTR: 1303

Homologous proteins: Top 10 BLAST Hits:

Sequences producing significant alignments:	Score (bits)	E value
CRA 103000001515936 /altid=gi 10946600 /def=ref NP_067277.1 hy... CRA 120000042903164 /altid=gi 13358640 /def=dbj BAB33045.1 (AB... CRA 87000000001314 /altid=gi 8923754 /def=ref NP_060871.1 gene... CRA 87000000001426 /altid=gi 7161864 /def=emb CAB76566.1 (AJ25... CRA 108000024647823 /altid=gi 12730486 /def=ref XP_003392.2 ge... CRA 18000005184360 /altid=gi 7505957 /def=pir T23688 hypothetical... CRA 18000005004115 /altid=gi 1730069 /def=sp P54644 KRAC_DICDI ... CRA 18000004912236 /altid=gi 464395 /def=sp P28178 PK2_DICDI PR... CRA 18000004910302 /altid=gi 1170689 /def=sp P42818 KPK1_ARATH ... CRA 89000000197925 /altid=gi 7295638 /def=gb AAF50945.1 (AE003...	760 733 556 554 423 335 217 203 202 201	0.0 0.0 e-157 e-156 e-117 1e-90 4e-55 8e-51 1e-50 2e-50

Docket No.: CL001078-DIV
Serial No.: TO BE ASSIGNED
Inventors: Ellen BEASLEY et al.
Title: ISOLATED HUMAN KINASE ...

EST:

Sequences producing significant alignments:
gi|13032240 /dataset=dbest /taxon=960...
gi|6588496 /dataset=dbest /taxon=9606 /...
gi|883123 /dataset=dbest /taxon=9606 /...
gi|946492 /dataset=dbest /taxon=9606 /...

Score (bits)	E value
1348	0.0
1021	0.0
702	0.0
236	1e-59

EXPRESSION INFORMATION FOR MODULATORY USE:
gi|13032240 prostate
gi|6588496 /lung
gi|883123 /whole brain
gi|946492 / Adult brain

Tissue Expression
whole brain

FIGURE 1B

1 MYAMKYMNKQ QCIERDEVRN VFRELEILOE IEHVFLVNLW YSFQDEEDMF
51 MVVDLGGD LRYHLQQNVQ FSEDTVRLYI CEMALADYL RGQHIIHRDV
101 KPDNILLDER GHAIKDFNI ATIICKDGERA TALAGTKPYM APEIFHSFVN
151 GGTGYSFEVD WWSVGVMAYE LLRGWRPYDI HSSNAVESLV QLFSTVSVQY
201 VPTWSKEMVA LLRKLLTVNP EHRLSSLQDV QAAPALAGVL WDHLSEKRVE
251 PGFVPNKGRL HCDPTFELEE MILESRPLHK KKKRALKNS RDNSRDSSQS
301 ENDYLODCLD AIQQDFVIFN REKLKRSQDL PREPLPAPES RDAAEPVEDE
351 AERSALPMCG PICPSAGSG
(SEQ ID NO: 2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

288-291 NKSR

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2

1 75-77 TVR
2 245-247 SEK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 4

1 42-45 SFQD
2 226-229 SLQD
3 298-301 SQSE
4 300-303 SEND

[4] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

152-157 GTGYSF

[5] PDOC00100 PS00108 PROTEIN_KINASE_ST
Serine/Threonine protein kinases active-site signature

95-107 IIHRDVKPDNILL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	185	205	0.727	Putative

FIGURE 2A

BLAST Alignment to Top Hit:

>CRA|103000001515936 /altid=gi|10946600 /def=ref|NP_067277.1|
hypothetical serine/threonine protein kinase [Mus
musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
/length=488
Length = 488

Score = 760 bits (1942), Expect = 0.0
Identities = 371/399 (92%), Positives = 381/399 (94%), Gaps = 1/399 (0%)
Frame = +1

Query: 103 VNFDFQILRAIGKGSFGKVCIVQKRDTCKMYAMKYMNNQQCIEDEVRNVFRELEILQE 282
Sbjct: 89 VNFDFQILRAIGKGSFGKVCIVQKRDTCKMYAMKYMNNQQCIEDEVRNVFRELEILQE 148

Query: 283 IEHVFLVNLWYSFQDEEDMFMVV DLLGGDLRYHLQQNVQFSEDTVRLYICEMALADYL 462
Sbjct: 149 IEHVFLVNLWYSFQDEEDMFMVV DLLGGDLRYHLQQNVQFSEDTVRLYICEMALADYL 208

Query: 463 RGQHIIHRDVKPDNILLDERGHAHL TDFNIATIICKDGERATALAGTKPYMAPEIFHSFVN 642
R QHIIHRDVKPDNILLDE+GHAHL TDFNIATIICKDGERATALAGTKPYMAPEIFHSFVN
Sbjct: 209 RSQHIIHRDVKPDNILLDEQGHAAHL TDFNIATIICKDGERATALAGTKPYMAPEIFHSFVN 268

Query: 643 GGTGYSFEVWWSVGVMAYELLRGWRPYDIHSSNAVESLVQLFSTVSVQYVPTWSKEMVA 822
GGTGYSFEVWWSVGVMAYELLRGWRPYDIHSSNAVESLVQLFSTVSVQYVPTWSKEMVA
Sbjct: 269 GGTGYSFEVWWSVGVMAYELLRGWRPYDIHSSNAVESLVQLFSTVSVQYVPTWSKEMVA 328

Query: 823 LLRKLLTVNPEHRLSSLQDVQAAPALAGVLWDHLSEKRVEPGFVPNKGRHLCDPTFELEE 1002
LLRKLLTVNPEH SSLQD+Q AP+LA VLWD LSEK+VEPGFVPNKGRHLCDPTFELEE
Sbjct: 329 LLRKLLTVNPEHRFSSLQDMQTAPSLAHVLWDDLSEKKVEPGFVPNKGRHLCDPTFELEE 388

Query: 1003 MILESRPLHKKKKRLAKNKSRDNSRDSSQSENDYLQDCLDAIQQDFVIFNREKLKRSQDL 1182
MILESRPLHKKKKRLAKNKSRD+SRDSSQSENDYLQDCLDAIQQDFVIFNREKLKRSQ+L
Sbjct: 389 MILESRPLHKKKKRLAKNKSRDSSRDSSQSENDYLQDCLDAIQQDFVIFNREKLKRSQEL 448

Query: 1183 PREPLPAPERSRDAEAEVED-EAERSALPMCGPICPSAGS 1296
EP P PE+ D + D EAE +ALPMCG ICPS+GS
Sbjct: 449 MSEPPPGPETSDMTDSTADEAFTPALPMCGSICPSSGS 487 (SEQ ID NO: 4)

>CRA|120000042903164 /altid=gi|13358640 /def=dbj|BAB33045.1|
(AB056389) hypothetical protein [Macaca fascicularis]
/org=Macaca fascicularis /taxon=9541 /dataset=nraa
/length=368
Length = 368

Score = 733 bits (1872), Expect = 0.0
Identities = 358/369 (97%), Positives = 361/369 (97%)
Frame = +1

Query: 193 MYAMKYMNNQQCIERDEVRNVFRELEILQEIEHVFLVNLWYSFQDEEDMFMVV DLLGGD 372
Sbjct: 1 MYAMKYMNNQQCIERDEVRNVFREL ILQEIEHVFLVNLWYSFQDEEDMFMVV DLLGGD 60

Query: 373 LRYHLQQNVQFSEDTVRLYICEMALADYLRGQHIIHRDVKPDNILLDERGHAHL TDFNI 552
LRYHLQQNVQFSEDTVRLYICEMALADYL GQHIIHRDVKPDNILLDERGHAHL TDFNI
Sbjct: 61 LRYHLQQNVQFSEDTVRLYICEMALADYLCGQHIIHRDVKPDNILLDERGHAHL TDFNI 120

Query: 553 ATIICKDGERATALAGTKPYMAPEIFHSFVNGGTGYSFEVDWWSVGVMAYELLRGWRPYDI 732
ATIICKDGERATALAGTKPYMAPEIFHSFVNGGTGYSFEVDWWS+GV MAYELLRGWRPYDI
Sbjct: 121 ATIICKDGERATALAGTKPYMAPEIFHSFVNGGTGYSFEVDWWSLGVMAYELLRGWRPYDI 180

Query: 733 HSSNAVESLVQLFSTVSVQYVPTWSKEMVALLRKLLTVNPEHRLSSLQDVQAAPALAGVL 912
HSSNAVESLVQLFSTVSVQYVPTWS+EMVALLRKLLTVNPEH SSLQDVQAAPALAGVL
Sbjct: 181 HSSNAVESLVQLFSTVSVQYVPTWSREMVALLRKLLTVNPEHRFSSLQDVQAAPALAGVL 240

Query: 913 WDHLSEKRVEPGFVPKGRLHCDPTFELEEMILESRLPHKKKKRLAKNKSRDNSRDSSQS 1092
W HLSEKRVE FVPNKGRLHCDPTFELEEMILESRLPHKKKKRLAKNKSRDNSRDSSQS
Sbjct: 241 WGHLSEKRVEPDFVPKGRLHCDPTFELEEMILESRLPHKKKKRLAKNKSRDNSRDSSQS 300

FIGURE 2B

Docket No.: CL001078-DIV
 Serial No.: TO BE ASSIGNED
 Inventors: Ellen BEASLEY et al.
 Title: ISOLATED HUMAN KINASE ...

Query: 1093 ENDYLQDCLDAIQQQDFVIFNREKLKRSQDLPREPLPAPESRDAEAEPMCG 1272
 ENDYLQDCLDAIQQQDFVIFNREKLKRSQDLPEPLPAPE RDAEAEPMCG E+SALPMCG
 Sbjct: 301 ENDYLQDCLDAIQQQDFVIFNREKLKRSQDLPEPLPAPEPRDAEAEPMCG 359

Query: 1273 PICPSAGSG 1299
 PICPSAGSG
 Sbjct: 360 PICPSAGSG 368 (SEQ ID NO: 5)

Hmmer search results (Pfam):

Scores for sequence family classification (score includes all domains):			
Model	Description	Score	E-value
PF00069	Eukaryotic protein kinase domain	197.7	1.8e-55
CE00359	E00359 bone_morphogenetic_protein_receptor	9.8	0.044
CE00022	CE00022 MAGUK_subfamily_d	9.2	0.013
CE00031	CE00031 VEGFR	1.6	1.3
CE00292	CE00292 PTK_membrane_span	-92.7	0.0016
CE00291	CE00291 PTK_fgf_receptor	-116.1	0.041
CE00287	CE00287 PTK_Eph_orphan_receptor	-125.0	2.8
CE00286	E00286 PTK_EGF_receptor	-131.8	0.0056
CE00290	CE00290 PTK_Trk_family	-206.7	0.58
CE00016	CE00016 GSK_glycogen_synthase_kinase	-248.2	0.0061

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
CE00022	1/1	86	107 ..	133	154 ..	9.2	0.013
CE00359	1/1	95	144 ..	272	326 ..	9.8	0.044
CE00031	1/1	79	171 ..	1051	1141 ..	1.6	1.3
CE00286	1/1	3	212 ..	1	263 ..	-131.8	0.0056
CE00287	1/1	7	212 ..	1	260 ..	-125.0	2.8
CE00290	1/1	4	215 ..	1	282 ..	-206.7	0.58
PF00069	1/1	2	228 ..	26	271 ..	197.7	1.8e-55
CE00292	1/1	5	233 ..	1	288 ..	-92.7	0.0016
CE00291	1/1	1	233 ..	1	285 ..	-116.1	0.041
CE00016	1/1	1	303 ..	1	433 ..	-248.2	0.0061

FIGURE 2C

1 GATGGAGCCC AGGCCACATG CTTGGGAAAG CGGGTGGTCC TGGGTGAGCC
51 TGTCTGCAG GAGGGTGAAGG TCCTGTCTGG AGTAGACACC CGGTCTGTGT
101 CCGCCGACA TGACAGGG CCCAGAGGG ACGGGAAGCC CAGGTGCCCT
151 GTCCCCCTGGT GAGTGGCTG TGGGGGTTGA GCACCCCAAG GGAAGGGTGT
201 TTCTTCAGA GATTCGGGG TAAGAGCTGA GCTGCATCGT GAGCAGGAGG
251 GGCAGAGGG GCCTGGAGAG GGTTGGTGTG CCTGGGCAAG CAGCTTCTG
301 GGGTGGTCA CAAGGACAG AAAGCCTGAG GCTTGGGACG GGCTGAAGAC
351 GGCCTGGC TGCGGGTCTG GGCAGCGCTC CAGGCAGAGG TGGGCTCTGC
401 AGCCCCCCC CCCCCGCCAG CTTCCCCCC AGCAGCAAGA GCCCCTGGCC
451 CAGCCAGTGG CCCAACCCA GGCTCTGTAG GGGCAGAGGG TGAGGGCCGG
501 ATTTTCACCA ATATTTGTCT CACAGCCTGT CTGGTCCAG CCCCAGGGCA
551 ACAAACAGCC TTCTGGAGC AGTTTCCAGA CCTGCAGTGG CCGCCTTGAG
601 CCTGCAGTGA CGCTCTGCAG GAGGCCGCGG GTGCTGGGC TGGCGCAGGA
651 AAGCACCGTT GCTTCTGGC CTGTGCAGAG TGAGGCTGGG GCTTCATCC
701 CGGGCACGGG ACTCTCGGC CTCTGGCGC CGTGTGCATG GGAGGAAGGC
751 CGTGTGGC AGCCACCGA CTCTGGCCG TTGGCAGTGG GAAGCGGAG
801 GAGGGGGTCC TGCCAGGGC AGCCAGGGGC TGCTGCAGCT TACGCTCACT
851 GTCATCTGA AACCTCAAC TGGCTTCAA ATAACAATT TAAAAAATGG
901 TCATAGAAA TGCAAGGAGT TCAGGAGAAA TCCCGCCCCC CCCCAGCCCTC
951 CCAAGGGC CTCGTCAAG CTTAGTCTCC GTCTGTTCTC GGGCTGCCCTG
1001 CAGCTGCC CGCTCTCTAG CAGGTGTGG CGCGTGTCA GGAGCCCCA
1051 TCAGCACACG CGCTCTCTGG CAGCCCCAA CACAAGGCTC TTCTGTCCC
1101 TTCAGGCTCA GCTTCCCCC CCCACCGGGC AGGGAGGTGC TGAGGGCAGC
1151 CCTGTTCTCA GCTTCTGGG GAGGCCATAC TTAGAGCCCA TTGGTCCAGG
1201 CCAAGGACT GTCTCCCGCA CCACAGTGT GCAGCCGGAT CCACCCAGGC
1251 CACCGCTTGG CACCATCAGA CATGCTTTCT TAGTTTGGC CAGCCGGTG
1301 CTCTGTGCA AGTGTGACTC CAAGTGGGGC CTTGCGGGAG GGAGGGTAC
1351 CGCTTTCTT CACAAAAGGA CTGAGTCTCA GGGAGGGTGC TCCAGTAAGG
1401 GCCCGGAGC CAGGATGTGA CTGAGACAGG TGCCCTCAGG GCCACACACT
1451 AACCACTACA GGAACCTCTG GGGGCAGAAAT CATGGCCTCC AAAACACCCA
1501 CATCAAAATC CCAGAACAAA CATGTTACCA TTCTGTGGAG AGAGGAATGA
1551 AGATTGCGA TTGAATTAG GTTGTAAATC ATCTGACTTA ATTTTTTTAG
1601 AGACAGGGTC TCGCCCTGTC ACCCAGGCTG AGTTCACTGG TACAGTCGTG
1651 GCTCACTGCA GCCTCAACCT CTCAGGCTCA AGCAATCTC CTTCTCAGT
1701 CTCTGAGTA GCTGGGACTA CGGGGTGTGTG CCACCCACGCT TTCCGGGCAT
1751 GCAGCCAGG GCCCAGGGCC ATCTGTGGCC CACCTTGAGA TCCAGAATCA
1801 TCCATTCTCT CCAGGGCCCC TGCTGGGCTC CAACTCTTG AGGACCAAG
1851 AGCAGAGGTT GTGGAAGGCC TTGGAAACGG GTCTGGATTA CCTGTCCCTGG
1901 GAAGGCTCT CCAACCTGA GTGTCAAGACA GGGGTTAGCT CTGCTGCTCA
1951 CAATTGTTG CTTTAATTCC TGGCTTCCCT TTGGGGATCT TCATCTCAA
2001 TTCTGATTGA CATCTCTGG CACAAGGGAC CCCCCCTGCTC ATTGTGCTT
2051 CTCACCGTC ACCTCACTCT CATCCTCACT GCTAAGCAAT TAGCCGTGT
2101 TTTGGGCT CATGTTGAC ACCGATGATC CATGCTCAGA GGGTACAGGC
2151 CTGAAGAGCT ATGTGGGGAC TGGGGCCCCC GAGGGGGTCC CGCTGTGGT
2201 GCAGCGGTTGCCCCCAAGCC CCACGCTCAC TCTGTGTGTC TCCTTGCGAG
2251 ACAGGGTGTGGGGCTGGC CTCACGGTGT TGAGACGGGA GTCGGTCTCT
2301 GGAGTGTGG GTGATGTGGC TCCAGGGTGA AGCTGAGCC ATGGTTGATG
2351 AGGCTTCTG AGGCAGGGGGC TCTGTGCTCA GTGGGTTTTC GCATTCTTC
2401 CCCAGTACCC CTCCGGGCTG CTTGCTGCTC CAAGCCCTGG AAAAGGATGT
2451 TGGGGTTAG GAAGGAAAAA CTCCATGCC AGGTCTGGC TGGTGAAGGG
2501 CGCTCCCCGC ACAACCCAGCC ACTGCTTGGC TCCAACCTCAC GCCTGGATGC
2551 TGTTAGGCTG GACCCCTGTG GTTTGCAGAT AGCCGCTGTT GACAGATGTG
2601 TCCTGTGCA GACTTGAAC GCAGGACTGA GTCTCAGGGA GGGGTCTTC
2651 CTAACAGCCA TGGAGCCAGG ATGTCACTGA GACAGGTGCC GCCAGGGCCA
2701 CACACTAACCA AGTACAGGAA CTCCTGGGG GCAGAACATC GCCTCCAAA
2751 ACACCCACAT CTTAATCCC AGAGCAAAATA TGTTACCATA TGTGGCAGAG
2801 AGGAATGCACT CGCCCGCGTC TAGGAGGGAG AGGGGCTGCG GCGCGTCCCC
2851 AGGTGTCCCC TTGTGTCTG ACGATGCTC CCAAGGCGTC CCCAGGATC
2901 TCCAGGTGTC CCCACGTGTC CCCAGGCGTC CTCAGGCGTC CTGCCGAACA
2951 GCCCTGTGCT CTCAGGTGT GCATTGTGCA GAAGCGGGAC ACGGAGAAGA
3001 TGTACCCCAT GAAGTACATG ACAACAGCAG AGTGCATCGA GCGCGACGAG
3051 GTCCGCAACG TCTTCCGGGA GCTGGAGATC CTGCAGGAGA TCGAGCACGT
3101 CTTCCTGGTG AACCTCTGTG GAGCCTGCA TGGCCTCGCT GCAGAAAGAC
3151 TTACCGTCTT GAAGCGGGGA AGGCACACTGG TATCCTCTCC TGCCCTGGT
3201 GCTCTTGGC CTCAGGTGT CGGAGATGCC TCTGCCACCC ACAGGGCTC
3251 TCTAGCCCTC CTACATCTATC CCCCCTCTG TGCCCCAGGC CTGGCAGTGG
3301 CCCAGGTGGC CATGACATGC TGGGGTTGGT TAATGCAGTG TCTCTTCTGA
3351 GCCTGCGGA AGACCAAGGGC TTCCCTACAA TGGAGATGTG CTCCCATGGA
3401 GTCTGTGCA CTAGTCAGAG AGGGAGAGAG TTAGGGACT GAAAAGCTA
3451 CCACTGTCACT CACCATCACC ATCACCAACCA TCATGCCAT CACCACCA
3501 ATCACCACTG CAACCATCAT CACTCTCATC ACCATCATCA CTATAATCAA
3551 CACAATCACT ATTGTCACTA CCATTACAC CACCAACAG ACCACAATCA
3601 CTGTTATCAC CATCACCAAC ACCCTCAGTC CTCACTACCG TCATCCTCAC
3651 CATCACCGTC ACCACCAACCA CCATCACTGC CATCGTCAAC ACCATGGATG

FIGURE 3A

3701 CTGGTTGTTA AATGCCAGCT CTTTGCCCAA CACTGTCAAG AGTGGTACCT
3751 ACACGGCCTC ATTTCTGTA ACAACCCCTCC GAGGCAAATG TCTGTATCCC
3801 CATTTACCG AAGAGGAGGC CGGGCAGCCT GAAGCACCCG GAGCTGGCAC
3851 TGTAGCTCTG CTCGCATTT GCCACTCCA GGTGCCTCTG GCCCCAGCTG
3901 GGCCACCTCC AGCACAGGGT GGTGTGTCTT TCCTCAGGAT CTGGGCTAG
3951 AGCTGCTCTG GGCTGGGTG CAATCAGTGC CTTGGGCAAG CCCCTCCCTCC
4001 TGGAATGCC TGGTGGCTGA TGCTGGGGTG GGGCTGTGGT CCTTAGGGGG
4051 AGTGTGTCAG CTGTGGAGC AGCCATGACT GGCTCCCCAG CTGTGCGCAC
4101 AACAGGCCCT CCATCGGTGC CCACAGGTAC TCCTTCCAGG ACGAGGAGGA
4151 CATGTTCATG GTCTGGACCC TGCTACTGGG CGGGGACCTG CGCTACCAC
4201 TGCGAGCAGA CGTGCACTTC TCCGAGGACA CGGTGAGGCT GTACATCTGC
4251 GAGATGGCAC TGGCTCTGGA CTACCTGCGC GGCCAGCAC A TCATCCACAG
4301 GTGTGTGCGT GGCAGACGGC GCAGGTACCT GCTGAGGTGG GCAGGGCTGA
4351 AGCAGCCTTA GGTCAGGCTG CGGGCACCGC GGCGTACTC CCTCAGAGCG
4401 GGTCTAGCTC CTCTGCCCA CCCTTGCTG AGTGCCTGCC CCCAGCTGTG
4451 GCACCTGTGC CGACAGGTCA AGCCCCATAG CTGTGTGCTT GGTGTCATC
4501 TGGGGGACCC TCGTCCCAGC AGCCCCAGCT GAGACTGGC ACAGTGGCT
4551 GTTAGGCTG GTGGACAGAC CACCAAGCTG GGTCACAGCA GGTGGGCTTC
4601 ACCTGGTCCA TTTAACTGAA GACTCCCTGTT TGCCCATCCA CCACATCCCA
4651 GGGAACTCAA ACTAATTTA ACATTAGCTT AAAGCAGATG AAATTAGGAA
4701 GCAGAGCTGG TGTGATGGCT CTGAAAATAA AATTTAAAAA AAGAAAATAG
4751 GAAGCAAGATT ATGAAGGAAG TGAAAATTGGG AAGCAGAAAT TAGGCTGAAA
4801 TCCGAGCAGA ATGAAACAAA ATGAAAATAT CTGTGAGGTA TATTTTAAAG
4851 TCGAATGGAC TTGGTTTGC ATTTCCTGTC TTGGGGACTC GGATGTCTGA
4901 TTATGACCTA GGCAACAGTC ACTGAGCACT GGCTGTGTAC CTGGAAAAGT
4951 TGGGACAAAG CAAGAGCCGA GGTGGCTTGG TCTCCTAGAG GCCGAGTCTT
5001 GGAGGGGGAG GGCAGACCTT GGCAGCAATT GCTCTCGTCC TCTGGGCTC
5051 CAGGGCCCTT CCCAGCATCT GGTGCAAGGT GTGTGCTGCT GCCCAGATGC
5101 CACAGGAAC GAAGTGGCTG ACTTCATCGC CTCTGCCCAAC ACGCAGGGT
5151 GTGAGGTCTT AGCATCATCC AAGGACCAAG TCAAGCTCCC AGGCCTCTGC
5201 CTCGAGTGGG TTGGTGGAT GTCTGGGGGA CTCCAGGGAT TGTGACAGAG
5251 ATTCCAGGGC AGAAACAGG CAGATTTCA ACTCACCTTC CCACTTCTG
5301 CTCTTTCTAG AGATGTCAAG CTCGACAAAC TCTCCTGGA TGAGAGAGGT
5351 GTGTGGGTTT GGGTGTGGC AGCCCAGGT GGTGGTGGCA GGGATGGGCC
5401 TGTCAAGGGGA GGAGGATCTC GCACGCAAGG ATGCATCTCT GTTCTGGGA
5451 CAGCCACACC TGACCCCTC CTGCACAGGA CATGCACACC TGACCGACTT
5501 CAACATTGCC ACCATCATCA AGGACGGGA GCGGGCGAGC GCATTAGCAG
5551 GCACCAAGCC GTACATGGGT GAGCCCAGC TGGGGTTCA GATGGGAGCT
5601 GGCTCTCTC AGGTGGGAAG GACAAGACCT CGGTGGCTTC TCTGTCCAC
5651 CCTGGAGGA GCCTGGCTC GGGATGTGGC CTCAAGGTGC CGGCCCCCTGTG
5701 CCCACGGTC CGGGCTGTA CCCCCGTGGCA GCTGTTTTC CTTCTTCTG
5751 TCGGAAAGCT CGGAGATCT TCCACTCTT TGTCAACGGC GGGACCGCT
5801 ACTCTTCAGA GGTGGAATGG TGTTCTGGGG GGGTGTGGC CTATGAGCTG
5851 CTGCGAGGAT GGGTATGGAC CCCCCCTGCAGC CCCCCGGCTT GGCTGCCAGG
5901 CCCCTGCTCT CTCGCCCCAC CAGTGTCTGG GAGGGGGTGG CTGCCCCAGT
5951 GCCCAGGTGC GCAGGGATGT CTCCACTGTG TCTGAGGAGT CACGCTTTTA
6001 TCGAATGTG TAGTTGGTA TGGAAATGCTT GAGCAGGAGG AGGAAGGACA
6051 GACTCACTGT GGTGCTCCCG GGGCGCTGCT GGTGCCTGCA GGCCAGCCTC
6101 TGTGGGGTTG GACAAGGCTG AGAACACTGGCC AGCAGGGGTG CTGCTCTGGA
6151 ACTTTCCACA AAAAGTTCT TTTGGGCC TGTGCTCTTA CCCTTGTGGC
6201 CACGGGAGG CCAGTCTGG AGACCGGGAG GCTGGGGGTC CTCTTGTGGA
6251 CCGTACCCCT CAGGGCTGCA CAGGACCCCCA CCTCTGAGGA AGCCAGCTCC
6301 CTCTGGGCC TCTGGGGCTG ATCTACCTGG ACCCAGGGCC CCTGGGATCC
6351 CAGCCAGATG GGGCGAGCAG CCAGGGCGCA GGACCCAGGC GTAAGCTTA
6401 TCTCACCCAG GCTCTCTGC GGCAGGTGGA GGCCAGGTG TGCTCAGAGC
6451 TGTGCTGCA CTGGGGTGG GGGGAGGGGG TCCTCTCAGG GCGATGGCAC
6501 CTGTGTCTGG CATTGTTCTG GGTGTCTCTGG GGGCCAGGGAG GACCTGCCA
6551 GCACTGCTCTC CCTGTCTCCA GAGGCCCTAT GACATCCACT CCAGCAACGC
6601 CGTGGAGTCC CTGGTGCAGC TGTTCAGCAC CGTGAGCGTC CAGTATGTCC
6651 CCACGTGGTC CAAGGAGATG GTGGCTTGC TGCGGAAGGT GAGCCCCAT
6701 CCCTAGGCTT CTCACCCCT CGAGCACCCA CCTCCCTCCC TCACTTACCT
6751 GCGGCTCGGG ACACCCCCCTC CAGTGCACAG TTAGTGCAGC TTCTGGCAG
6801 GCACAGATCC CTTCACTGCA ACCTGTGGGG GCCTCCGCAG ATGGCAGCCC
6851 CAAGCCCCAG GAAGCGAGCT GGTGGCAGGG TCTGTGGCCC TCTCATGGCG
6901 AGCCCTACCA GGTCACTGGT GTCTCTGGTA TCCCCCTGAGC TGCGTCTCCA
6951 GGCACCCCTCA CAGCAGGCTT GTGCGCCCCG CCGCCGGTC ACCACAGAGG
7001 AGCCCTGAGC CACGTCCTCT TGTCCACCGT ATTGGAAGGA GGGTGTGGGG
7051 TGGGTGGGTT CAAGCCCCATG CTATCTCCGG GACCCCTTGC CCCAGTGCCT
7101 CCTGGGGAAAG GTGGGCAAGC ACTGCCCACT GCAAACACCT CTCAGGGAC
7151 ATCAGCTGG CAGGGACAG GGCAGGGGTG TGGCCATCAG TAGTGCCTCC
7201 CATTGTTGAT CTGGTCTAG GCTGGGGCTG TGCTTGTACT GGGCTGCACA
7251 GCTCCCTGTT CCGAGCCTCG CATTAAACACC ATAGGGGGTT CGGAGTCAGA
7301 GCCAGGCCA GGCAGGACAG GGAGGGAGT GAGTGTGCCA CACGGGGCCG
7351 GCTGCCCTCC GGCCCCCTG TCTCAGGAGC GTGGGGCTC CTGCCCTGGA

FIGURE 3B

7401 ATTGTAGCCA AGCAGCCTAA AGCCTTGGGG AGGCCTTGCC TGCCGGGCC
7451 TCTCCCAGC CCCGAGAGTC TCTTAACCT GCTGTAGCCC CATGAAGCTC
7501 AGTCACACCT GCCCAGGTGG CTCACAAGGT GGCACTGGC TAGAGAGGC
7551 CTGCGTGGG ACTGGGGATG ACCCCACACGC CCAAGCCAG GTCTGGGAAA
7601 CCTCGCACGG GGTCTGGTC TGCGGCATT TCCCTGGAAA GGCGGGAGGT
7651 GCCAGGCTG GGATGTTGCT TCCCAGGCCA TGCATGGCTG CCCCGGGCTC
7701 ATCTGGCTG TGGAGGTCCC ATGATTGGT GAAGGAAGTG GCTCTGGGAT
7751 AGTTACTGTG AGGGCAGCA TGTGCCAGT GTTAGCCCT AGCCGGGCCT
7801 CGGCTGCCAC CTCTGGCAA ATCCCAGCA AGCCTTCCT GCAGATCCCT
7851 CTGCTGTCT CTGCGCAGG GGGTTAGGT AGCAGCACTG AGAACAGGCG
7901 TCCCTGGGC CACATGCTG GCCAGGCCAG GTGCTTGGC TGATGTGGC
7951 CGTCGGCACCC ACCTTCTC GCGTGGCCCT GAGGTTCTG AATTCTGAAC
8001 CTGAGGCTTG GTGGGACCC CTCAGGTG CCCTGGCCCT GGGGTTGGCGG
8051 GCTATTCCTG GCTGTGGGC TGTGGGCCCT GGACCCCTCTG ACTCATGCCT
8101 GGTTGCACTG CCTCACTGTG AACCCCGAGC ACCGGCTCTC CAGCCTCCAG
8151 GACGTGCAAG CAGCCCCGGC GCTGGCCGGC GTGCTTGGG ACCACCTGAG
8201 CGAGAAGAGG GTGGAGCCGG GCTTCGTGCC CAACGTAAGC CTGTGGCGG
8251 CTCAGGTGGG GGGCCCTGGG GATGGATGT GCGTCCTCCA CGGGCCGGGG
8301 CTAGCACCCT ATCCCTGTG AGAAAGGGCC TCTGCACTGC GACCCCACCT
8351 TTGAGCTGGA GGAGATGATC CTGGAGTCCA GGCCCCCTGCA CAAAGAAG
8401 AAGCGCTGG CCAAGAACAA GTCCCGGGAC AACAGCAGG AGCAGCTCCA
8451 GTCCGTGAGT GCCAGGGCAG GCTCAGGGCG CGGCGGGGG CTGGGCTTGG
8501 GGCTCTCTC TACACCGAG CAAGGTGTG GGGGACCCCT GACAGTGCAC
8551 ACGCTCTCGA ATGGCAGCAG ACCGTTTCTC GAAGTCCTGA GAAGGCCAGA
8601 GACCTCTTCTT CTGCTTCTC CAGCCCCCAC CTCGCTCTT ATGAAGCAGG
8651 TGGGCAGGGA CAACCAGGGC TGGGGTTATG AGTGCACGGG GATGGCCATG
8701 TGAAGGCTTC GTGCTTGGCC AGGTGTGCTG GTGTTGGTT TGTGTGGGG
8751 GACGGCTATG TGAAGGCCCT ACACTCGCCC AGGTGCGTGC GCATCAGGTA
8801 TGTGTGGGG GACAGCCATG TGAAGCCCTC ACACTCACCC AGGTGCGTGC
8851 GCATCAGTTG TGTGTGGGG GACGGCCATG TGAAGCCCTC ACACTCGCCC
8901 AGGTGTGCTG GCTTTGGT TGTGTGCGAG GATGGCCACA TGAAGGCCCTC
8951 ACACTCGCCC AGGTGCGTCA GCATCAGGTG TGTGTGGGG GACGGCCATG
9001 TGAAGGCCCTC ACACTCGCCC AGGTGCGTTC ATGTTGTG TGCAAGGGATG
9051 GCCATGTGAA GCCCTCACAC TCACCCAGGT GCGTTGATGT CAGTTGTG
9101 TGCAGGAGCA GCCATGTGAA GCCCTCACAG TAGCCCAAGGT GTGTCGGTGT
9151 CAGTTGTG TGTTGGGATG GCCACGTGAA GCCCTCACAC TTGCCCAGGT
9201 GCGTTGATAT TAGTTGTG TGCAAGGGATG GCCACGTGAA GCCCTCACAC
9251 TCACCCAGGT GCGTTGATGT CAGTTGTGAG TGTGCGCAGG GATGGCCACA
9301 TGAAGGCCCTC AGACTCGCCC AGGTGCGTGC GCTTTGGTTG TGTGTGCAGG
9351 GACGGCCATG TGAAGGCCCTC ACACTCGCCC AGGTGCGTCA GCATCAGTTG
9401 TGTGTGGGG GACGGCCATG TGAAGCCCTC ACACTCACCC AGGTGCGTGC
9451 ACATCAGTTG TGTGTGGGG GACGGCCATC TGAAGCCCTC ACACTCACCC
9501 AGGTGTGCG GTGTCAGTTG TGTTGTGCGGG GATGGCCACG TGAAGGCCCTC
9551 ACACTCGCCC AGGTGCGTGC ATATTAGTT TGTGCGCAGG GATGGCCACG
9601 TGAAGGCCCTC ACACTCACCC AGGTGCGTTC ATGTCAGTTG TGAGTGTG
9651 CAGGGATGGC CACGTGAAGC CCTCAGACTA GCCCAGGTGT GCTGGCTTG
9701 GTTGTGTG TGAGGGGGC CATGTGAAGC CCTCACACTC GCCCAGGTGC
9751 GTCAGCATCA GTTGTGTG TGTTGGGACG CATGTGAAGC CCTCACACTC
9801 GCCCAGGTGC GTCAAGCATCA GTTGTGTG TGTTGGATGG CACGTGAAGC
9851 CCTCAGACTA GCCCAGGTGC GTCGGCATCA GTTGTGTG CGGGGACAGC
9901 CACGTGAAGC CCTCACACTC GCCCAGGTGT GCGGGCTTTG GTTGTGTG
9951 CGGGGACGGC CACGTGAAGC CCTCATGCTC ACTCAGGGCAT GCTGGTATT
10001 TGGGCTGCC AGGACAGGTG ACCACGAATC AGGTGGTTGA AGAACAGCAA
10051 TCGCTCTCTC TGAGAGGATC TGAGTCGTAA TGAAATGGTC TCCTTCACAG
10101 CCGGCTGTG GTAACTACT CTGTCTCTG CAGCTCCCT GTCTTGATAA
10151 TTGGCTGTG AGGACGGGG TAAGGTGAAC CCCTTGGGCA GTTATGTG
10201 GATCTCAGTT TCTGTAACAA GGAAGTCCAG GCATGGTGCA GCTCTGTTCC
10251 CTGCTTGGG GTCTCACAG ATATGTGAGCT AACATTGAGG TCGTGGCCTT
10301 GTCAGGTGCA GCTCTGTTC CTGCTTCTG GTCTCCCCAG AGTGTGAGCT
10351 AATATTGTG GAGGTGTG TCTCATCAGG GATTGTGACAG GTGCTGTGGT
10401 TGAAATGTT CCTTTAAAC TGCGTGTGGA ATTTGTTTCC TATTGTGATG
10451 GTGGTAGAGA TGGGACTTTT GGGGGCTGAT GGGGCCACGT AGGTTCTCC
10501 AGCATGGATG GGGTTAACATGC TGTTGTAGAA GGGTGACTTT AGTCCTCTT
10551 TGAGTCTTTG ATCCTCTGCT ATGTGAGGAC GTGGGTGTTCC CAATGTGGAC
10601 GTGGTTCTG TTCCATGTGA ATGTGATATT CACAATAGAC CATCAACAGG
10651 CTCCCTTTA ATCAGCAGAT TTAAAAAGAA ATGTGTTGTC TCATGGCTTG
10701 GAGGCCTGAG TCCAAAGTTA AGATGTCAAG AAAGCCGTGC CCCCTCTGAA
10751 GGCTCTCGG GGAGGAAAC CAGTCCTTG CCCCCTACCC TCCGGTAGAG
10801 GCTGCCCTGG CCTAGACGCC TCCCCCCCAGC CCCTGCTTCC CTGCCCGCTG
10851 GGGTCGGCT GTGTGTGCGT CTCCCATCTCC TCCCCCTTCC TCATAAGGAC
10901 ACCAGGCATT GGATTTAGGG CCCACCCCTGA TCCAGTATGG CCCCATCTTA
10951 TCTTGATGAT ATCTGCAAG ACCTCACTTC CAAATGAGGT CACATTCA
11001 GGTACCCAGG ATTAGAATT TGAGTGTGCA TTTTTGGGA CACAGTTGG
11051 CCCATACAC CAGGATGTGG CTGATATTCA CCAAGGAGTA GCTATGGTTG

FIGURE 3C

11101 TGTGTTGATG TCAAGGGTGA GGTGATGACC CTGGGTCCCT CGGTGGTCCC
11151 CTTGCCCTG AGCTGCTG AGCCTGTT GGATGTCCTG GGAAACTCTT
11201 GTGCCTCAGC CCCCGTGAG CCTCCTCAGA CCTGGTGGC CCTGTGTTGC
11251 TCCGGGAG AAGACGGGT TCAGTCCCC CTCACCATG ATGTGGGGG
11301 CAGGGGGGG GTCATGCCCT GGGTGCCCT ATTTTGGGG GAACACGGCC
11351 CCCCCAGGG TCAGGCTCCC ATCCCTCGCC CTCCTCCAGG ACGGCTGCC
11401 GCAGCCCCGG GTGCTCTAG GCAGAATTGC TGGTGGAGAG CTGCTGCTG
11451 CCAGGGGCC ACTGTGAGGC ACTGCTGAGA GCCACAGGAT GGTTGGAAGG
11501 TTCTCGGGGT TGGGGGTTCT TTGGCATTGC CCCCATTGGA TGTTTAAGTT
11551 TTCCCTACCA GAGCATGTTA AGAGGCAAGG CTCTGGGTG TAGAACACAGG
11601 CCCAGGATGA GTAGGAACC CTATGGGG ACTCAGGGAT GGACAGTGTG
11651 CAGGGCCAG CTGGCCATGC TGAGTTCCA GGAGGCTCTG GCTGGGAACA
11701 GTTAAGGCCA GGCACCTGTG AGCGGGAGGA GCTCGGCTT GTCTGGGTT
11751 GGTTGTTGG AGATGTTTG GCTTGAGGGT AGGAGGTTGTT CTGAAAGGAA
11801 AGCATCACTC CAAAAAAA GTCCCCTGTAACCTTGAG GCTGAGAGAG
11851 GTTTTTGAA ACAGCTTTAT TTTGATATAA TTCACATTCC ATGCAATATA
11901 CAGTGATCC ATGAAAGCA TATAATTCCA TGGTTTTAA TATAGTCACA
11951 GGGCTGCA TTCTCCACCA CAATCTGATT TTAGAACCTT TTCACTGAAAT
12001 GTAAGGAAA GACCCCACTT ATTAGCAGTC ATGCCCCATT CCCCTCTTCT
12051 CCCCTCCCT GGGCACCAAG AAGCTACTTT CCGTCTCTG AGGGTTGCC
12101 GTTGTTGGCG TTTCATGGAA GTGGAGTTAC ACACATATGTG GTCTTGCAG
12151 CTGGCTCTT TCACCTCGCA GGATGCTTT GAGGCCCTC CACGTTGTAG
12201 CCTGTCAGT CTTCATTCAG GTTGTATGGCT GAGTAATATT CCACATATGT
12251 ATCACCTTC CTTTATCCAG TCATCAGTT ATGAGTATT GTTCTTTCCA
12301 CTTGTTAACCA TTTTTCATTA TCATGAATAA TGCTGCTGT AACATTACA
12351 TACAAGCTT TGTTGTTA TGTTACTTTA TTTTTGGGG GCACATACCT
12401 AGGCATGAAC CCGCTGGTC ATATGTGACT CTGTCCTCA TGTTGGAGG
12451 AAACACCTAC CTTTCTTAC AGCAGGTGTG CCATTTATG CCCCTACCA
12501 CAGTGTGTA GGGTTCTAAT TTCTCCATAT ATTTACCAAG TCCTGTTATT
12551 GTCTGTTTA TTTTTTTAA AATCATAGTT ATCTTAGTGT GCAGTGATTG
12601 TGTTGTTATG GTTGCATT CTCTGATGAT ATTGAACATC TTTTGAGGTG
12651 TTTTATCAGG CATTGTTGAG AGAGAAATGT CTATCAAAT GTTTTAAAAT
12701 TTTTATTGTC TTTTTTTAG TCTACTCTGA CAATATATT TAATTGGCAT
12751 ATTTATTTA CTTTATTTT TTTTAGAGAC AGGGTCTTGC TTTATTACCA
12801 AAACCTGGAGT GCAGTGTGAA ATAGAAGGCT CACTTCAGCC TTGACTTCC
12851 GGGCTCAAGT GATCCTCCCT GCCCCCAGTG CCAGAATGGC TGGAACTGTG
12901 GGTGTCACC ACCACACCTG GCTCATTGA AAAAATTGTT GTTGTAGAGA
12951 CAGGGTCTCA CTATGTTGC CAGGTTGGTC TCAAACCTT GGAGTCTCC
13001 CACCTCAGC TCCAAAATG CTGGGATTAC AGATGTGAGC CACTGTGCC
13051 GACCTAATTG GTGTATTTA GACCATTCA GATTTAAAGCG ACCAGGGAGG
13101 CTGAGCAAG AGGACTGCTT GAGTCCAGGA GTTGGAGACC AGCCTGGCA
13151 ACAAGGTGAG ACCCATCTGT ATTAGTCTAT TTTCACACTG CTAATAAAGG
13201 CACATCTGAG TCTGGTAAAT TTATACAGGA AAAAGGTTTA ATGGATTTC
13251 AGTCCACAT GGCCTGGGAG GCCTCACAGT CGTGGAGGC AAGGAGGAGA
13301 AAGTCACATC TTACATGGAT GGCGGCAGGC AAAGAGAGAG CTTGTTCA
13351 GAAACTTTG TTTTAAAC CATCGGATCT CATGAGACTC ATTCACATATC
13401 ATGACAACAG CACAGGAAG ACCCCGCCCC ATAATTCAAT CACCTCC
13451 CAGGTTCTC CCACACATG TGGGAATTGT GGGAGTCACA ATTCAAGCTG
13501 AGATGGGAT GGGGACAGAG CCAAACCATA TCATTCTGCC CCAGCCCC
13551 CAAATCTCA TGTCTCACA TTTCAGAAC AATCATGCC TCCCAACAGT
13601 CCCCCATAGT CTTATTTGG CATTAACTCA AAAGTCCACA GTCCAATGTC
13651 TCATCTGAGA CAAGGTAAAGT CCCTTCTGCC TATGAGTCTG TAAAATCAA
13701 AGCAAGTGC TTCTAGATA CAATAGGGGT ACAGGCTTG GGTAAATTCA
13751 GCCATTCAA ATGGGAGAAA TTGGCCAAA CAAAGGGCT ACAGGCC
13801 TGCAAGCTG AAATCCAGCA GGCGCTGCAA ATCTTAAAGC TTCAAAATGA
13851 ACATCTTGA CTCTATCTC CACATCCAGG TCATGCTGAT GCAAGAGTTG
13901 GGTTCCCATG GTCTTAGGA GCTCTGCCCT TGTGGCTCTG CAGAGTACAG
13951 CCTTCCTCCC GGCTGTTTC GTGGGCTGGC ATTGAGTGTG TGTGGCTTT
14001 CCAGGTGCTG GTGCAAGCT GTTGGTGGAT ATACCATTTCT GGGGTGTAGA
14051 GGATGGGGC CCTCTCTCA GAGCTCCACT AGGCAGTGCC CCAGTGGG
14101 CTTTGTATAG GGGCAACAC CCCACATTTC CTTCTGCAT TGCCCTAGCA
14151 GAGGTTCTCC ATGAGGGCCC CACCCCTGCA ACAAACCTCT GCCTGGACAT
14201 CCAGGTGTT CCATACATCC TCTGAAATGC AGGCAGAGGC TCCCAAAACCT
14251 CAATTCCTGA CCTCTGTGCA CTCGGGGCT CAACACCCACA TGGAAGCTG
14301 CACAGCTGG GGCTTGTACCT CTCTGAAGCC ACAGCCTGAG CTGTACCTT
14351 GCCCCTTCA GTCATGGTG GAGCAGCTGG GATGCACTG AGCAAGTCCC
14401 TAGACTGCAC ACAGCAGAGG GACCCCTGGAC CTCGCCCATG AAACCATTTT
14451 TTCCCTCTAG GCCTCTGAGT CTGTGATGGT AGGGGCTGCC GCAAAGGTCT
14501 GTGGCATGCC CTGGAGACAT TTTCCCCATT GTCTGGTGA TTAATATTCA
14551 GTTCCTTGTGTT GCTTATGCAA ATTTCTCTC AGAAAATGGG GTTTCTTT
14601 TTTTCTCTCT TTTTTTTTT TTTTGAGAC AGTCTTGCTC TGTCACCCAG
14651 GCTGGAGTGC AGTGGTGCAA TGGCGGTCA TTGCCACTGC AACCTCCGCT
14701 TCCTAAGTTC AAGTGTATTCT CCTGTCTCAG CCTCCCAAGT AGCTGGGATT
14751 ACAGGCAAGC ACCACACAC CCAGCTAATT TTTGTATTT TAGTAGAGAA

FIGURE 3D

14801 GGGTTTCACC ATGTTGGCCA GGCTGGTCTT GAACTCCTGA CCTCAGGTGA
14851 TCTGCCGCC TTAACCTCCC AAAGTGCTAG GATTACAGGC GTGAGCCACC
14901 GTGCCAGCC AGGAGTTCT TTTCTATTGC ATTGTCAAGGT TGCAAATTTT
14951 TTGAACTTT ATGCTGTTTC CTTTTAAAAA TGGATGCCT TAAACAGCAC
15001 CCAAGTCACC TCTTGAATGC TTTGCTGCTT AGAAATTCT TCTGCCACAT
15051 ACCCTAACAT ATCTCTAAA TTCAGAGTTC CACAAATCTC TAGGGCAGGG
15101 GCAAATGCT GCCAGTCTCT TTGCTTAAAG CATAACAAGA GCCACCTTG
15151 CTGAGTCC CAACAAGTTC CTCATCTCCA TCTGAGACCA ACTCAGCTG
15201 GACTTCATTG TCCATATCAT TATCAGCATT TTGGTCAAAG CCATTCAACA
15251 AGTCTTAGG AAGTCCAAA CTTTCCCCACA TTTTCCCTGTC TTCTCTGAG
15301 CCCTCCAGAT GTTCCAGGC TCTGCCATT ACCCAGTTCT AAAAAGTTGC
15351 TTCCACATT TCAGGTATCA TTTCAGCAGC GCCCTACTTT ACTGGTACCA
15401 ATTTACTGTA TTAGTCTGTT CTCACGCTGC TAATAAGAC ATATCCGAGA
15451 CTGGGAAATT TATAACGAA AAAGGTTAA TGAGCTTACA GTTCCACATG
15501 GCTGGGGAGG CCTCACAAATC ATGGCGGAAG GCAAGGAGGA GCAAGTCACA
15551 TCTTACATGG ATGGCAGAGA GAGCTTGTGC AGGGAAACTT TTGTTTTAA
15601 AACCATCAGA TCTCATGAGA CTCATTCACT ATCATGACAA CAGCACAGGA
15651 AAGACCGCC CCCATAATT AATCACCTCC CACTGGGTTTC CTCCCATGAC
15701 ACACGGAAAT TGTTGGAGTC ACAATTCAAG CTGAGATTGG GGTGGGGAGA
15751 CAGCAAAACC TTATCACCG CTCTATAAAA GACAAAAAAA TTAGGCAGGC
15801 ATAACAGTGC ATGCCCTGAG TTCCAGTGTGAT GTGAGAGGAT TGCTTGAGTC
15851 CAGGGAGTTG AGACAGGCT GGGCAACATG GCGAGACCCCT GTCTCTACAA
15901 AAAAAGGTTA TCTGGGTGTT GTGGGATACA CCTGTGATAC CAGCTACGCA
15951 GGAGGCTGAG GCAGTAGGAT TGCTTGGGCC CAGGAGTTCA AGGCTGCAGT
16001 AAGCTATGAT CATGCCCTG CACTCCAGTC TGGGTAACAG AGAGACACGC
16051 TGTCCTGTA ATAATAAAGT GGTCAATTAT ATAGTTCAAT ATGATATCTA
16101 CTTTATTTGT AACTGTAGTC TATTATTTGG TCTTCTTTT TCCCTATTTT
16151 TCTGCCCTT CTGGTTAA TTAAGCATT TATATTATTC TAGTTTATCT
16201 CCTCTCCTGG CCTGTTAATT ATACTCTTT TTGAAATATT TTAGTGGTT
16251 GGCTCTGACA TTGCACTATA CTTTACCAT ACAGTCTACC TTCACTGAC
16301 ACTCTGCCCT CTCATGTCA GTGGGATGCC TTGTGACGGC ACCTCTCGT
16351 CAATCCTCT GTTCCCTGATG ACATTGCTGT CATTCACTTC ATTATCTGT
16401 ATGCTATAAT TGCTCATTAC ATTGTACTA CTGTTATTTT AAACAGTTAT
16451 CTTTTGGATC AATTAAAGAA AATTAAAAAT TTCACTTAC CTCTATTCT
16501 TCCCTCTCTA AAGTCTTCC TTTCTTATG CAGACCCAAAG TTGCTGACCT
16551 AAATCATTTC CTTTCCCCCT GAGGAACCTTC GTTTAACATG TCTTATAGGA
16601 CAGGCTCAAC AGAGATGAAT TCCCTTCCCTT TTGTTTGTG CAAAAAAGTC
16651 TTTACTTCA CTTTAAAGA ATAATTTCAC TGGATATAGA ATTCTAGATT
16701 GGAGCTTTT TACTTTCAA CACTTTAAAT ATTCACTTCC GCTCTCTTCT
16751 TGCTTATGTG ATTCTAAAC GGCAGTCTGC TCTAATTCTT TTTCTGTAAG
16801 TAGAATTCT CGCCACCCCC CACCCCCCAGC TCATTTCAAG ATTTCTCTG
16851 TCTTTGGTT TCTGCAATT TGAATAATGAT ATGCCCTAGGT ACAGATTTTT
16901 TTTAAATATT CATTCTGCTT GGTGTTCTCT GAGCTTCTTA GATCTGTGGT
16951 TTGGTGTCTG TCTTAAATT CAGAAAATTTC CAGTCATTA TTATTCAAA
17001 TATTCTGTC TCTTCTCTT CTCCCTGCTGG AAATCCAATT ATGTGTGTG
17051 GGTGCCGTTT GAAATTGTCC CACAGCTCTT GGATATTCTG TTCACTTTT
17101 TCACTCTTTT TTTTTCTCT TTGCAATTCA GTTCGAGAAG TTTCTGTTGA
17151 CATTCTCTCA AGCTCATAGA TCCCTCCTCA TCTGGATCCA GTCTACTGAG
17201 GAACCATCAA AGGCATCTCTT ATTCTGTGTTA CAGTGTTC TACTTCCAGC
17251 ATTTCTCTTT GATCCCTCTT TGGAGTTCC ATCTCTCTAC TTACATTGCC
17301 CTCTCTGTTT CCATCTGTC TTGCTGCTT TTCCCCCTGGA GCCCTTGCCA
17351 TATTAATCAC AGTTATTCA AATTCTCTGT TTGATAATTTC CAACATGGGT
17401 GCCATATCTG GATCTGTTCT AATGTGTGCA TTGCTCTTCA ACACTCTATT
17451 TTTTCTCTTA AGACCTGTA GTCTGCCCTT TAATGTTGT TGAAAGCTGG
17501 ACGTGATGTA TCTGGATAA AGAGCTGAAG TAGATGGGCC TTTAGTGTGA
17551 GGTTTATGTA ATCTGGCAG GTTGGGGCA GGGTTAAAG TCTGCTGTAG
17601 CTGTGGGTAC CAGAGGCTTC ACATTGTTT TCATTTCCCG GGTTGCCCTT
17651 GGGCTTGCTT AAATCCTCTT CCTCAGAGAG AGTCTGCGTC TTGTGCCCT
17701 CTCTGCTGGA ATCCCTGTC AACTGCGGAG GCCCTGTGGG TTGTTGGGA
17751 AGATGGGGG GAGGGAACTC TTCCACAGTC TGTGACCAAA TCTCAGTCTT
17801 GGGGGCTGT GCCCTTCAC AGTTGTTGAT CTGCTTTTCC CCTCCCCCTTA
17851 GGTGAGACAG GCTAAAGCGG GGTACAGGCT GAAAAAACAGT CCTCCCCCAA
17901 GTGAGATAAG CTTTCTCTT GGAGAGCAAA TTCCATTGTC TGTGGAGAAT
17951 GCTCTGGGTG TATTTCACAG TGGTACTGT CCCCATCCCA TGCCAGAGCC
18001 AGGAAGGGAT CATCCTGGC TTCAATTAGA CCTGGCAGGG TTCCCTGGAGG
18051 GGAAACCCAC AAAACGTTGG GGGCCTCATAGA AGACCGCAGC CGCAGGAGTT
18101 CTGCAACACGG CCCCGGCCGC TCCCCAGAGC TGCCCAAGGAG GTGTTCCCGC
18151 ACACCATCAG TTCTGCTCCA GGCCAGCAGG TCTCAGCTGT GACTTTGCTC
18201 ACCTGTCTCT CCAGACTTGG GGGTCGCCGC AGCGCTTGA CCTCAGCTCT
18251 CTGATGGTC CAGGAAAGT AATTGATTTT CATTAGTTC GAATTTTCC
18301 TGTCGAGGG ATGGGAGTGA GGCTTCCAC GTTGTGTTGCC GCAGAGCAGA
18351 AGCCAGAACT CTTGCTCGT CTCTGTTT CTTCGTCGGA GTCATGTCTT
18401 TTATTTCTG AACCGAGGT TGTGGTCAGA CAGGAGATTG GCAGGTATTT
18451 CCTCCTGTC CATGGCTAT CTTTCTCTT TCTTGAGGGC ATCATTGCA

FIGURE 3E

18501	GCAGGGGAAGT	CTTGGACCTT	GACGTTGTCC	GATGGGCTG	TTTCCTCCTT
18551	GGCTGCTCTG	CTTTGCTAA	GGATCCATCA	CCTCAGCTGA	GGCCACAGGG
18601	ATTACTTCCC	ACACTTCTT	CTAAGCTTGT	TATAGTTTC	GGCTTCTCGT
18651	TTAGCTGTG	GATTCTTT	GGGTTAATT	TTATGAAACA	TGTCAAGGTG
18701	GGGTCCAGCT	TCATTCCTC	ATCGGTTGGAT	ATCTGACTGT	CCTAGCACCA
18751	TTTGCTGAAG	AGAGGATTCT	TTCCCCATTG	AATTGCTTT	GACATCATAC
18801	CTTGTTTTTT	GACTTGGCCT	TTTATCCAT	TGGTCAGAC	TGTCGCCCCG
18851	CGCCGGGACT	GCACCATTT	GATAACTGTA	GCTTTGTAGC	AAGCTTCAA
18901	ATCAGGACT	GTGAATCCC	CAGTTTTGTT	CTTTTGAC	ATTATTGAC
18951	TATTCTGGGT	CCCTGGATT	TCCACTGAAT	GCTGAGGGGT	CTGACAGTTG
19001	CATCTGAGCT	GCCAAGCAGG	TTTGTGGGG	TGCTAGGGAC	TAAGGCTGC
19051	TCCATTTCCTC	AGGCCCTCC	TCGCTGTGGG	TGACATCTGG	GGTCGAGGCG
19101	TGTGTCAG	CATGTGTTAA	GGTGCACAGG	GTGCCCTGAG	ATGGGGATT
19151	CTGGTCCAGT	TACTCAGAAA	GTGCATCCAG	GAGAGACCCC	CGCCCTTCTC
19201	GGGATGGGGG	ATGCGAGCAG	AGCTTGGCTT	TCAAGCAGAA	ATCTGGAAAC
19251	CCTGTGGGG	GTGGCTCGG	ACTTTCAAGGG	ACCTGGAGCC	TCACTTGTGG
19301	TTCAAGAGGT	CCCTGCCCTG	AGGGAGCTGG	GGCTCTAAC	ACCCACACCA
19351	GTCACTCATG	GGCGAAGGGC	CCCTCTGTGG	CCTTGTGACT	GTGTGTGTC
19401	TGGCAAAAGGG	TTCCAGGCC	CCAAGGAGGA	GGCAGGGGCT	TGCAAGGAGA
19451	GAGCAGGGCA	GATCCAGGCC	AGGCACAGAC	CCTGGCATGG	GGTACTGCTG
19501	GGCGCGCTAC	ACACACCTTG	GGAAAGGGGT	CCCTGTGAA	AGGGGCTCTG
19551	GTCAACACATC	TAGGTGACAC	AGCCCCGCTT	GGGGCCTGCT	CAGAGCCACC
19601	CCTTCAGAGT	GGTTCTGGAG	CAGCTCTTCA	GGCTCTCTGGT	GGCCTCTCTG
19651	CCTAGGAAAA	CATGGCTGTG	GACHTTGCAG	GATGACCAAC	AGCCCCTGCC
19701	ACTGGGCTGC	ACACAGGGCC	ACGACGGGGC	CTCATGTTCT	ACATCACTGG
19751	CGCCCCACCCC	AGCCCCCTCC	ACCTTGTGTC	GCTGTGAATC	GCAGGATCCC
19801	AGCGGCTCA	TCGGACCTTC	ATTCTGTAGT	AGTCTGAGCC	TTAGGTCAAC
19851	GTCACTCTCT	CAGGGCGGGC	CGAGTGTGCA	GACTTGTCTG	TCTATATCAG
19901	GGTAGACCA	TAAGGACAGG	TCCCCTGCCA	GGATCGTGAC	TCTTTAGGG
19951	TCTTGATGAC	TGACAGGGCA	AAGCTTGGAG	GGAGAGAGCC	ACATGGAGAG
20001	GAGGGCACAC	CCCGCAGAGG	GCGTGGGAGT	CTGCCGGCTT	CTTCTCGAGT
20051	GAGGGGCTG	GGAGTGTGCTG	CTCACACCGT	TCAGCTGTGG	GGTGTGGGAA
20101	CCTTGGCAAG	GACACCACCA	GGAGGAGCGG	GCGGGACAGT	CCCACACTCG
20151	GGACCAAGGT	AGAGAGTCCA	CATTCCCGCT	GCCAGGGTGA	AACCCCTCAC
20201	GGCTAGGAAG	TCCAGGAGAG	ACACGGGGG	CACTGCTCTG	GGTGTGGGAA
20251	CCACACAGA	GCCACTTGGC	AATCAGCATT	ATCTGGGTT	ACAGGGTGTG
20301	CGTAGCTGC	TCCCCAACCC	AAAGAAATGT	TTGAGGCCCT	ACAGGGGCAAG
20351	CCTCGTGGGA	AGAGGCCAA	GGTGAGACGG	CTTGTCTTT	ATCTGGGGG
20401	GCACGTCGA	GTGTGACGCA	CGTCAGCAA	TCCCTGAACC	GTTCATCGGA
20451	GAACAGCCTC	CTGCACTCTC	CAACTCTGT	TCGTGGGTT	ACAGGGTGTG
20501	CAGAGTACTT	GCCACTTGGC	AATCAGCATT	AACTAGATCC	ACAGGGGCAAG
20551	GCATGGTGC	TCACGCTGT	AACTCTGAGA	CTTTGGGAGG	CCGAAGCAGG
20601	CGGAGCACCT	GAGGTCAAGG	GTTCAAGACT	GGCGTGGCCA	ACATGGTGAAG
20651	ACCCCGTCTC	TACAAAAATT	AGCTTGGTAT	GGTGTGGG	ACCTGTAATC
20701	CCAGCTACTT	GGGAGGCTGA	GGCAGCAGAA	TCACTTGAAC	CTGGGAGGCG
20751	GAAGTGGCAG	TAGGCAAGA	TCACGCCACT	GCACTCTAGT	CTGGGTGACA
20801	GAGCGAGACC	AAAAATAGGA	TCCATGGATA	GCAGGCAAGA	GTGTCCAGGT
20851	GTTCGAGGCA	CAGACGACAC	TGTGACAGGG	AAAGATCCCC	TTAGGCTCTG
20901	CTGGGGCGGT	GGAAAGCATGC	TGTTGTCCTG	CTCGGGTGA	CGCAGACTGT
20951	TGTCAAGCAT	TGCACTAACG	TGCTTCCCTC	ACTGGCAAT	CGCATGGGGG
21001	GGTGGTCCCG	TAAGATGGTA	ACACTGGGTT	TTGCTGTAGC	TTTTGTATGT
21051	CTAGATAGGG	TTGAGCTTTC	TGGTGTGTAC	CCACTCACAC	GTCCCCCTCCG
21101	ACCTTCAAG	CCCCAGCTCC	TCCCCTCCAG	GGCCTTGGCT	GTGACGGGGG
21151	TGACTTCTA	TGGATCTGAG	GTTCTGTGTT	CCTCACAAAGT	GGGCATCTCTC
21201	TGGCCTCAGC	TGCAAGGAGGT	GGGGGCCCTT	TTATGCCAC	CCGAGGCTG
21251	CACTCTCTG	CACTTTTCA	TGTGACTTGG	CTCATCTGGG	TCTGTCTT
21301	GCTCACCGT	TGGTAGTGA	CAACGTACC	ATCCAAGTT	ACGGTCACCA
21351	TAATGATGCT	TCCCCAACAC	CATGCCAACG	TCAGGCCGGC	AGCACCCCC
21401	CCACCCCCACC	CACGCCCTCC	CCCGACCCCT	GGCTCTGGG	AAGTGGCTCT
21451	GTCTGCTGAA	GGAAAGTAGTC	CTGCTGCC	ATCACACACCC	AGTAAGGGTG
21501	GGTCTCTGCCA	GGGGCAGCGG	CGCTCCCAA	GCTTGGCTG	AGGACCTGCT
21551	TCTAAGAGAC	CCCTGGTCCC	AGGATTCTT	GGGGCAGGGC	CCAGAACAGCAG
21601	GCCTGGGACA	GGTGTGTG	TGCTGTGATG	AGGGGCTGGG	AGAACCCGGT
21651	ATGAGACGGG	AAAGGCCCGG	CAAGGGAGTG	GTTCCTTCA	GAAGTCCCCA
21701	GAGAACAGCT	TCTGCTGTTG	CTCTGAGGCC	CCACGGAGCA	AGTCCAAGCC
21751	ATCCACCCAG	AGGCAAGGGG	GCTGGGGCTT	GGCATCTCG	GGCTTGGGGT
21801	AGTCACCCGC	AGAGATGCGA	GCTCCCCGGG	CAGTCTGGCT	GCTGGAGGGC
21851	CGGGGCACTT	CAAATAGCCC	AGAGGCCCTG	ATCCAAAGCC	ACAGGTGGAG
21901	GGCCGATGGG	GATGCCAGCA	CACTCACTG	AGGGGACATG	GGGGGAACCT
21951	GGACAGCGTC	CCCCAGCTC	ACGTGTGCT	TTCCATCCAC	AGGAGAATGA
22001	CTATCTCAA	GACTGCTCTG	ATGCCATCCA	GCAAGACTTC	GTGATTTTTA
22051	ACAGAGAAAA	GTGAGTGTG	TGGGGTGTGG	TGGGGTGTGG	GGCAGAGGAGG
22101	AGGAAAATGG	GGCTAAGGT	AAAGTTCTT	TGGGCCAC	AGGGGGCACC
22151	TGTGGGCTG	GGGTGCTG	CCCTGCTCTC	TTGGGGACT	GTGAGCAGCA

FIGURE 3F

22201 GCTATGGAGG GGAGCGGCGG GAGGCCCTG CCAGGCTCTG GCATGTTGT
22251 GCTCCACCGG GGGCCCTG TGGAATCTT CTGGGGAGAG ACACATCATT
22301 TGCCCAGATG AGGGGGTGTG ACTTCCTAGG AGGCCCCATC AGAGCCACGT
22351 CAACTCCCC ACCCAGGAC GCCCTCAGTC TCTCAGCAGA CCTTCTCTGA
22401 ATGTCAAGGC CCAGGGACAA GAAAGGGCAC AGATGACTGG CAGCAGGCAA
22451 GGCAGGCCAG AAATAGCAGC AGCTGCCAC GTGGGGCCCA AGGGAGGATG
22501 GATGCTCCCT CTGGCCGAC GGGGCAAGGA GGGCTCTG GAGGAGGTGG
22551 GTCTGAGCTC TTATGGACAG GACGTGCAAG GAGCAGCTG CAGACGGCTG
22601 AGGGCACTGA CTGGCACCT GGGGATCAGA CGACCGGGTG AAGAATGAGG
22651 CTTAGCCGAG CCTCATTCAG AAGTCACTGA CCTATGGCAC CTGCACAGTC
22701 AGGCCTTCG GCTTCTGGT GGAAACATGC CGAGCCTCGC CAGCATGCTC
22751 ACGTGGCCCC ACCCGTCCCC AGGCTCCCTG CAGTGTGTC GGGAGCATGG
22801 CCTCTCCAGC AGACACCGAG CCTGTGGCCC ACGTTGGGC ATCCACGCCA
22851 TGGCTCATCC CATGAGCCCG TGGGGCAGGTG ATGGGACCGT GAGGCAGGG
22901 AGGTGGGGGC ATAAGCCCT CCATGTGTT CGGCCACCC AGGCTGAAGA
22951 GGAGCCAGGA CCTCCCGAGG GAGCCTCTCC CGGCCCTGA GTCCAGGGAT
23001 GCTCGGGAGC CTGTGGAGGA CGAGGCGGAA CGCTCCGCC TGCCCATGTG
23051 CGGGCCCTT TGCCCTCGG CGGGGAGCGG CTAGGCCCCG ACGCCCGTGG
23101 TCCTCACCCC TTGAGCTGT TTGGAGACTC GGCTGCCAGA GGGAGGGCCA
23151 TGGGCCAGG CTTGGCATTG ACGTTCCAC AGCCTGGC TGGCGGTGCG
23201 CACAGTGCCT CGGACACATT TCACACCTCA GGCTCGTGGT GGTGCAAGGG
23251 ACAAGAGGT GTGGGTGAG GGGACACCTG TGGAGGGCAT TTCCCGTGGG
23301 CCCCCGAGAC CCCCTAGAT GGAGGAAGCC CTGCTGGGC CCCTCTTAC
23351 GCTCACGGGG AGCTGGGGCC ATGGATGGGA CAGGAGTCTT TGCCCTGCT
23401 CAGCCCGAG GCTGTGCAAG GCCCTCGTCA CAAGGTGACC CTTGCAAGCAC
23451 AGGCCGGGG TGCCCCAGGC TCAGCTCAAGT TCTTGGAGGT CAAGGGCATG
23501 GTTGGGTA GTGGGTGGG AGGTGAATGT TTCTAGAGA TTCAAACATGC
23551 TCCAGCAATT TCTGTATACT TTTCACCTCT GAGAATTACA ATGTGAGAAC
23601 CGCTCGATGT TGCTATGTT GCGTACGTCC TGTGTGTCG TGGCGTCA
23651 GCGGTGCTT GCGTTTCTG GTTGGCCTGG ACTTGGGGCA GCCAGTGGGG
23701 TGGGCAGCTC CTCAAGGGCAG AGCTCCCGGA CCATGGCTTT GGGGTGGGTG
23751 CCTGCTCTGG TGCCCCCTGGA GCGCTAAGGC TGTTGGAGGC AGAGACGGTC
23801 CTGGAGGAGC AGGAGCCAG GACAGCACCG TGACCGTGG AGCCGCCGCA
23851 GTGCCGGCA GTGTTGGCC CTCCATAAAAG GGACGTATCC CTCTCACTGT
23901 GGCTGGTGG TTCTGTGTTT GGAAGTGTAA CTAACTGGGT AAACGCCCTG
23951 TGTCTCTC TCTGGTCTCG TGAGGAGGAG ACAGGCTCAG CCCGTCAGCC
24001 CAGCGCTCCA GACAGGCCCTG TGCTGGTTTC CTCTGAGGAA ATGGGTGTGG
24051 CGGGTCTGTG CCCCTTCCCA GGACAGCGGC CATAGTGGAC ATGTGCCCTAG
24101 ACCTGTGTC ATGAGCCCCA CTGCACCCCT GGCAAAACAGG GCCCTCCCGT
24151 CCTGGCTGG GCTGCGAGAT GGAGATGACA ACGGCCAAAG AACATTGGG
24201 GAAGAACCGG CCATGCCAGC AGCAGAGTCA GAAGTCCGAG GGGATAGAAC
24251 GCAGCTTCCC GTCCCCCACC CACCCCTGTC CTCTGAATCA TGGCAGAAC
24301 TAGCCTTCA GCCCTCAGCA GCTCACATGG GGGACACGGC ACCCAAATCA
24351 CCACCAAGAA GGGTGGCCCG GTCTCTGGCA GGGCCCCAGAG GCGCCGTGTA
24401 CTCGGTGGAA GTCTGGCAGT GTCAAGAGACA GGCTCGGGC AAGGACAGGT
24451 GTGGGGTTT GAATAAGTGC ATTGGGGAA CATGGCAGGG TGGTGCACCT
24501 TGTCTCTCTT CGAGACATG TGAGGAGTGTG GTGCTGTCT GTTCCCTTG
24551 ATCGCCCCCC ACATGGGGC AGAGTGGGG AGTCTGGTGT GGGGGACATC
24601 AGCTCCCCA TCTGGGCCAG AGGGAGGCCCG GGGAAAGGAAA TGCTGAGGC
24651 CCAGGGCCTT CGCCTGGGAT CTGCACAGCT TTATAAGCAG CCCAGGGTGA
24701 GAGATGGGGC TGCTCTGTG TCCCAGAGAC CACGGCAGGA ATTCTCTGT
24751 CACCATCGGT GCATGGGAG GGGCCAGAGA ACCGGTGGCA CAAGGTGTCC
24801 CTGGCTCTCT GCTCAACAAA CAGCGAGTGC CAGTGAATG CGACCAAGGCC
24851 CCGCTCTGG GATGAGGACA ACCGTCTGGG AACGTCCACG CACCTTTAT
24901 GAGACACAGC ACGCGCAGC ACCGCAGTCAC CACACCGGG GCTCGGGTCA
24951 GCCTCATAGC TGCCCGGGTG TGAGTGTG TGCTGCTGT GTGAGCAGCG
25001 CCCACCTGGG TGCCGGTGGT GTGTGCTTCA CTTCACACA CTCCCGTGA
25051 TGCTCCCCGG CCTTCTGGGG TGGCTGGGG TGTCGGTGA GAATGTAGGC
25101 GGGGGGGGGGG GGGGGCCTCT GTCCCGCCCT GTATGTTGGC TGCCCTCTGC
25151 CCCGGCTGG ATGGTGGCTC CCTCTGGGGC CATCTTCCCT GTGAGAGGGT
25201 GCGCTCTCC CTGGCATTGA GGGGAGAAGA GCTCGGGCTG CAGGAGTCGT
25251 CGGACAGCC CACAGCCAGC CAGGCCCGC CTGTGAGAG ACGGCGTGGG
25301 GGAGAGGAGA CGGGGCCCTTC CTTCCATGCA CAGGCGGGTT CAAACCCAGA
25351 CGTCTTAAAT GGGCCTGATT CACATCAGAG CGAGGATGAC TGCTGTCCA
25401 GGCCTGGGT GGCATGCACA GTTCTGGC TACAGTGTCC TCAGTGTACA
25451 AAGTGTCTAC TAAGAAGCT ACGGAAATAC ACAATGTGA ATAAGAGGAC
25501 AGTGTCTTCC TAAAGGATCG CAAACCTCC CTGGATGAGG GCTACATGGA
25551 AGCTTAGGTG TGGCCTGGG GGTGCGTAAA AGGGACCTC CACGGCGGG
25601 GCT
(SEQ ID NO: 3)

FEATURES:

Start: 3000
Exon: 3000-3118

FIGURE 3G

Intron: 3119-4126
 Exon: 4127-4300
 Intron: 4301-5310
 Exon: 5311-5348
 Intron: 5349-5478
 Exon: 5479-5568
 Intron: 5569-5758
 Exon: 5759-5862
 Intron: 5863-6571
 Exon: 6572-6688
 Intron: 6689-8108
 Exon: 8109-8234
 Intron: 8235-8322
 Exon: 8323-8454
 Intron: 8455-21992
 Exon: 21993-22060
 Intron: 22061-22942
 Exon: 22943-23081
 Stop: 23082

SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
1931	G	A	Beyond ORF(5')			
4232	G	A	Exon	75	T	T
4309	G	T	Intron			
4622	G	A	Intron			
4708	T	C	Intron			
5787	C	T	Exon	150	N	N
5884	C	T	Intron			
8481	C	T	Intron			
8754	G	A	Intron			
8847	G	A	Intron			
11114	G	A	Intron			
19741	A	G	Intron			
20908	C	T	Intron			
22728	T	C	Intron			
23406	C	T	Beyond ORF(3')			
24078	G	A	Beyond ORF(3')			
24777	G	A	Beyond ORF(3')			

Context:

DNA Position

1931 AGTCAGTGGTACAGTCGTGGCTACTGCAGCCTCACCTCTCAGGCTCAAGCAATCCTC
 CCTTCTCAGTCCTCTGAGTAGCTGGACTACGGGTGTGTGCCACCGCTTTCCGGCAT
 GCAGCCAGGAGGCCAGGGCCTCTGGGCCACCTTGAGATCCAGAATCATCCATTCT
 CCAGGGCCCTGCTGGGCTCAACTCCTTGAGGACAGAGAGCAGAGGTTGTGAAGGCC
 TTGAAACGGGTCTGGATTACCTGCTCTGGAGGTCTCTCCAACCTGAGTGTAGACA
 [G,A]
 GGTTAGCTCTGCTGCTACAATTTGTGCCTTAATTCTGGCTTCCCTTGGGATCTT
 CATCCTCAATTCTGATTGACATCCTTGGCCACAAGGGACCCCCCTGCTCATGATGCTC
 TCAACCGTCACCTCACTCTCATCCTACTGCTAAGCAATTAGCCGTGTGTTGGCATC
 AGTGTGACACCGATGATCATGCTCAGAGGGTACAGGCTGAAGAGCTATGTGAGGACT
 GGCGCCCCGGAGGGGTCCTCGCTGTGGTGCAGCGGTGGCCCCAACGCTCACT
 4232 CCTCAGGATCTGGCTCAGAGCTGCTCTGGCTGGGGTCAATCAGTGCCTGGCAGGC
 CCCTCTCCTGGGAATGCTCTGGCTGATGCTGGGGCTGTGGCTTAGGGGGA
 GTGTGTCAGCTGTGGGAGCAGCCATGACTGGCTCCCCAGCTGTGCCACACAGGCC
 CATCGGTGCCAACAGTACTCCTTCAGGACGAGGAGGACATGTTCATGGCTGGACCT
 GCTACTGGGCGGGGACCTGCGTACACCTGCGCAGCAGAACGTCAGTTCTCGAGGACAC
 [G,A]
 GTGAGGCTGTACATTCGAGATGGCACTGGCTCTGGACTACCTGCGCGGCCAGCACATC
 ATCCACAGGTGTGCGTGGCAGACGGCGCAGGTACCTGCTGAGGTGGGGGGCTGAAG
 CAGCCTTAGGTCAAGCTGCCGGCACGGGGCGTACTCCCTCAGAGCGGGCTAGCTCCT
 CTGCCCCACCCCTGCTGAGTGCCTGCCCTCAGCTGTGGCACCTGTGCCACAGGTCA
 CCCCATAGCTGTGCTGGTGTCCATCTGGGGGACCTCGTCCCAGCAGCCCCAGCTGA

FIGURE 3H

4309 CCTGGTGGCTATGCTGGGGTGGGGCTGGTCCTTAGGGGAGTGTTCAGCTGTGGGA
GCAGCCATGACTGGCTCCCAAGCTGCGACAACAGGCCATCGGTGCCACAGGT
ACTCCTTCAGGACGGAGGACATGTTCATGGTCGACCTGCTACTGGCGGGACC
TGGCCTACACCTCAGCAGAACGTGCAATTCTCCAGGACACGGTGAGGCTTACATCT
GCGAGATGGCACTGGCTGGACTACCTGCGCGGCCAGCACATCATCCACAGGTGTGC
[G, T]
TGGCAGACGGCGCAGGTACCTGCTGAGGTGGCGGGCTGAAGCAGCCTAGGTCAAGGCT
GCCGGCACGGCGCGTACTCCCTCAGAGCGGGCTAGCTCTGCCCCACCCCTGCCT
GAGTGCCTGCCCGAGCTGTGGCACCTGTGCCACAGGTCAAGCCCCATAGCTGTGTGCC
TGGTGTCCATCTGGGGGACCTCGTCCCAGCAGCCCCAGCTGAGACTGGCACAGTGGG
TGTAGCCCTGGTGACAGACCACCAAGGCTGGTACAGCAGGTGCCCTCACCTGGTCC
4622 CAGGTACCTGCTGAGGTGGCGGGCTGAAGCAGCCTAGGTCAAGGCTGCCGGCACGGCG
GCCGTAECTCCCTCAGAGCGGGCTAGCTCTCTGCCCCACCCCTGCCTGAGTGCCTGCCT
CCAGCTGTGGCGACAGGTCAAGGCTAGCTGTGGCACCTGTGCCACAGGTCAAGCCCCATCT
GGGGGAGCTCGTCCCAGCAGCCCCAGCTGAGACTGGCACAGTGGCTGTAGCCCTGG
TGGACAGACCAACCAAGGCTGGTACAGCAGGTGCCCTCACCTGGTCCATTAACTGAAG
[G, A]
CTCCTGTTGCCCATCCACACATCCCAGGAATCAAACATAATTAAACATTAGCTAA
AGCAGATGAAATTAGGAAGCAGAGCTGGTGTGATGGCTCTGAAAATAAATTAAAAAAA
AAAATAGGAAGCAGATTATGAAGGAAGTGAATTGGAAGCAGAAAATTAGGCTGAAATT
CCGAGCAATGAAACAAATGAAAATCTGTGAGGTATATTAAAGTCAATGGACTG
GTGTTGCATTCTGCTTGGGACTGGATGTCGATTATGACCTAGGACCAAGTCAAC
4708 CTCCCTGCCCAACCTTGCTGAGTGCCTGCCCTGCCCCAGCTGTGGCACCTGTGCCACAG
GTCAAGCCCCATAGCTGTGCTCTGGTGTCCATCTGGGGGACCTCGTCCCAGCAGCCCC
GCTGAGACTGGCACAGTGGCTGTGGCACCTGGTGGACGACCAACAGGCTGGTCA
GCAGGTGGCTTCACTGGTCCATTAACTGAAGACTCTGTTGCCATCCACACATC
CAAGGAATCAAACATAATTAAACATTAGCTAAAGCAGATGAAATTAGGAAGCAGAGC
[T, C]
GGTGTGATGGCTCTGAAAATAAATTAAAAAAAGAAAATAGGAAGCAGATTATGAAGGA
AGTGAATTGGGAAGCAGAAAATTAGGCTGAAATTCCGCAATGAAACAAATGAAAAT
ATCTGTGAGGTATATTAAAGTCAATGGACTGGTGTGGTCACTGGCTGTGACCTGGAAA
TCGGATGTGATTATGACCTAGGCACCAAGTCAGTGACACTGGCTGTGACCTGGAAA
GTGGGACAAGCAAGAGCCGAGGTGGTGGTCTCTAGAGGCCAGTCTGGAGGGGG
5787 CACCTGACCGACTTCAACATTGCCACCATCATCAAGGACGGGGAGCAGGCCATTA
GCAGGCACCAAGCCGACATGGGTGAGCCGAGCTGGGTTCCAGATGGGAGCTGGCTTC
CTCCAGGTGGAGGACAAGACCTCGGTGGCTCTGTCCCACCCCTGGAGGCCAGCTGG
TCTCGGGATGTGGCTCAAGGTGGCCCTGTGCCACGGTCCGGCTGTGACCCCGT
GGCAGCTGTTTCTCTTCTGTGCGAAAGCTCCGAGATCTTCACTCTTGTCAA
[C, T]
GGCAGGACCGGCTACTCCTTGAGGTGACTGGTGGCTGGGGGTGATGGCTATGAG
CTGCTGGAGGATGGGTATGGACCCCTGAGGCCCCGGCTGGCTGCCAGGGCCCTGC
TCTCTGCCCCACCAAGTGTGGGAGGGGGTGGCTGCCAGTGCAGGTGGCAGGG
TGTCTCACTGTGCTGAGGAGTCACGTTTATGAAAGTGTGAGTGGTATGGAATG
CTGAGCAGGAGGAGGAAGCAGACTCACTGTGGTTCCGGGCGCTGTGGCCT
5884 GTTCCAGATGGAGCTGGCTTCCAGTGGGAAGGACAAGACCTCGGTGGCTTCT
GTCCACCCCTGGAGGAGCCTGGTCTCGGATGTGGCTCAAGGTGCCGGCCCTGTGCC
ACGGGTCGGGCTGTGACCCCGTGCAGCTTTTCTCTGTGAAAGCTCCG
GAGATCTTCACTTTGTCAACGGGGACCGGCTACTCTTGAGGTGGACTGGTGG
TCGGTGGGGTATGGCTATGAGCTGCTGCGAGGATGGTATGGACCCCTGAGCCCC
[C, T]
GGCTTGGCTGCCAGGCCCCCTGCTCTGCCCTCAGCTGCTGGGGAGGGGGTGGCTGC
CCAGTCCCAGGTGCGAGGGATGTCTTCACTGTGCTGAGGAGTCACGTTTATCGA
AGTGTGAGTTGGTATGAAAGTGCCTGAGCAGGAGGAGGAAGCAGACTCACTGTGGTT
TCCGGGGCCGCTGTGGTGCCTGAGGCCAGCTGTGGGGTGGACAAGGTGAGAA
CTGGCCAGAGGGTGTGCCCTCGAACAAAAAGTTCTTGGGGCCCTGTGGCCT
8481 GTGCTGTGGGACCACTGAGCAGAAGAGGGTGGAGGCCGGCTCGTGCCAACGTAAGC
CTGTTGGCGGCTCAGGTGGGGGCCCTGGGATGGATGTGGCTCTCCACGGGCCGGGG
CTCAGCACCCATCCCTGTAGAAAGGGCTCTGCACTGCGACCCACCTTGAGCTGG
GGAGATGATCTGGAGTCCAGGGCCCTGCACAAGAAGAAGCAGGCCAGGGCAAGAACAA
GTCCCCGGACACAGCAGGGACAGCTCCAGTCCGTGAGTGCAGGCCAGGGCTCAGGGCG
[C, T]
GGCGGGGGCTGGCTGGCTTGGCTCTTACCAACCGAGCAAGGTGTGAGGGACCCCTG
ACAGTGACACGCTCGGAAGTCCAGCAGACCGTTCTGAAGTCTGAGAAGGCCAGAG
ACCTCCCTCTGCCCTCCAGCCCCACCTCGCTCTTATGAGCAGGTGGCAGGGAC
AACCAAGGGCTGGGTTATGAGTGCACGGGATGGCATGTAAGCCTCGTGTGGCCA
GGTGTGCTGGTGTGGTGTGCGGGACGGCTATGTGAAGCCCTCACACTGCCA
8754 CGTGAATGCCAGGGCAGGCTCAGGGCGGGCGGGCTGGCTGGGCTCTCTCT
CACCGAGCAAGGTGTGCGGGACCCCTGACAGTGCACAGTCTCGAAGTCCAGCAGACC

FIGURE 3I

	<pre> GTTTCTGAACTCTGAGAAGGCCAGAGACCTCCCTCTGCCTTCCCAGCCCCCACCTC GCTCCTTATGAAGCAGGTGGGCAGGGACAACCAGGGCTGGGTTATGAGTCACGGGAT GGCATGTGAAGCCTCGTCTGCCAGGTGTGCTGGTGTGGTGTGCGGGGAC [G, A] GCTATGTGAAGCCTCACACTGCCAGGTGCCTGCCAGGTGCAGGTATGTGTGCCGGGACA GCCATGTGAAGCCTCACACTGCCAGGTGCCTGCCAGGTGCAGTTGTGTGGGACG GCCATGTGAAGCCTCACACTGCCAGGTGCCTGCCAGGTGCAGTTGTGTGGGACG GCCATGTGAAGCCTCACACTGCCAGGTGCCTGCCAGGTGCAGTTGTGTGGGACG GCCATGTGAAGCCTCACACTGCCAGGTGCCTGCCAGGTGCAGTTGTGTGGGACG </pre>
8847	<pre> GCACACGTCTCGGAAGTCCAGCAGACGTTCTGAAGTCTGAGAAGGCCAGAGACCTC CTTCTGCCCTCCCAGCCCCCACCTGCCCTTATGAAGCAGGTGGGCAGGGACAACCA GGGCTGGGTTATGAGTCACGGGATGCCATGTGAAGCCTCGTCTGCCAGGTGT GCTGGTGTGGTGTGTGCGGGACGGCTATGTGAAGCCTCACACTGCCAGGTGC GTCGGCATCAGGTATGTGTGCCGGACAGCCATGTGAAGCCTCACACTCACCCAGGTGC [G, A] TCGGCATCAGTTGTGTGTGCGGGACGGCATGTGAAGCCTCACACTGCCAGGTGTG CTGGCATCAGTTGTGTGTGCGGGACGGCATGTGAAGCCTCACACTGCCAGGTGC TCAGCATCAGGTGTGTGCCGGACGGCATGTGAAGCCTCACACTGCCAGGTGC TTGATGTTGTGTGCAGGGATGCCATGTGAAGCCTCACACTCACCCAGGTGC TGTGAGTTGTGTGCAGGGACAGCCATGTGAAGCCTCAGACTAGCCAGGTGTGCG </pre>
11114	<pre> AGACGCATCCCCCAGCCCCCTGCTTGCCTGCCGTGGGCTGGCTGTGTGCGTCTC CATCTCCTCCCCCTTCTCATAAAGGACACCAGCATGGATTAGGGCCACCCCTGATCC AGTATGGCCCCATTTATCTTGATGATATCTGCAAAGACCTCACTCCAAATGAGGTCA ATTACACAGTAACCCAGGATTAGAATTGAGTGTGTCACTGGGACACAGTTGGCCC ATACCAACAGGATGTGGTGTGATATTACCAAGGAGTAGCTATGGTTGTGTTGATGTC [G, A] GGTGACGGTGTGACCCCTGGGTCCCTCGGTGGTCCCTGAGTCTGCCGTGAGCC TGTGGTGGATGCTCTGGGAAACTCTTGTGCTCAGCCCCGTGCAGCCCTCTGACCTG GTGGGCTCTGTGTTGCTCTGGGAGAACAGCAGGGTGTCACTGCCCTCTCACCAGTGT GGGGGGCAGGGTGGGTGATGCCCTGGGTGCCCTGATTTGGGGGAAACAGGGCCCCC AGTGGGTGAGGCTCCATCTGCCCTCTGCCAGGACGGCTGCCAGGCCCTGGGTGT </pre>
19741	<pre> GTCAGAGGAAGAGCAGGGCAGATCCAGGCCAGGCACAGACCCGCCATGGGTACTGCTG GGCGCTACACACACCTGGGAAGGGACTCCCTGTTGGAAAGGGGTGTTGTCACACATC TAGGTGACACAGCCGGCTGGCGCTGCTAGAGCACCCTTCAGATGGTCTGGAG CAGCTCTCAGGCTCTGGCTCTCTGCCCTAGAAAACATGGCTGTGGACGTTGCA GATGACCAACAGCCCTGCCACTGGCTGCACACAGGGCACGACGGCGCTATGTTCT [A, G] CATCACTGGGCCACCCCAAGCCCTCCACCTTGTGCGTGTGAATCGCAGGATCCCA GGGCTCAGTCGGACCCCTATTCTGAGTAGTGTGAGCTTAGGTACCGTCACCTTCTC AGGCCGCCAGTTGCAACTGTTGCTATATCAGGGTTAGACAGAGAGATGCTGA GACACAGCAGATCACCCAGCTGCTCTTGTGATGACTAAGGACAGGTCCCCCTGCCAG GATGTCAGTCTTGTGGGAGGCAAGTGCAGGGCAAGCCTGGAGGAGAAGCCA </pre>
20908	<pre> CTGAGGTCAAGAGTTCAAGACTGGCTGGCCAAACATGGTGAACACCCGTCTTACAAAA ATTAGCTGGGTATGGTGGGACCTGTAATCCCACTGTTGGAGGCTGAGGAGCA GAATCACTGAACCTGGGAGGCGGAGTTGGCAGTGCAGCAAGATCACCACTGCACT AGTCTGGGTGACAGAGCGAGACCAAAATAGGATCATGGATAGCAGGAAGAGTGTCCA GGTGTGAGGCAAGACGACACTGTGACAGGGAAAGACTCCCTAGCCCTGGCTGGG [C, T] GTAAAGCATGCTGTGTCGCTCGGGTGAACGCAACTGTTGTCACGCTTGCATAAAC GATGCTTCCGTCAGTGGCAATCGCATGGGGGGTGTCCCAGTGGTAACACTGGG TTTGTGTCAGTTGTATGTCAGTAGGGTTGAGCGTTCTGGTGTGACCCACTCAC ACGCTCCCTCCGACCTTCAGAGCCAGCTCCCTCCAGGGCTTGGCTGTGACGTG GGTACCTCTATGGATCTGAGGTTGTGTCAGTCAACAGTGGGATCTGGCTCA </pre>
22728	<pre> CACAGATGACTGGCAGCAGGCCAGAAATAGCAGCAGCTGCCACGGTGGG CCAGGGAGGATGGATGCTCCCTCTGCCAGGGCAAGGGAGGGCTCTGGAGGAGG TGGGCTGAGCTTATGGACAGGACGTGCAAGGGCAGCACGTGCAAGCGCTGAGGGC TGACTGGCACCTGGGATCAGACGACGGGTGAAGATGAGGCTAGCCAGGCTCATT CCAAGTCAGTGCACCTGACAGTCAGGCCCTCGGCTTGTGCTGAAACA [T, C] GCCAGCCTGCCAGCATGCTCACGTCCCCCACCGTCCCCAGGCTCCGCCAGTGTG TCGGGAGCATGGCTCTCAGCAGACACCGAGCCTGCCCCACGTTGGCATCCACG CATGGCCTATCCCATGAGCCGTGGGAGGTGATGGGACCGTGAGGCCAGGGAGGTGGG GCATAACGCCCTCCATGTGTTCTGCCACCCAGGCTGAAGAGGAGCCAGGACCTCCGA GGGAGCCTCTCCCGCCCTGAGTCCAGGGATGCTGCCAGGCTGTGGAGGACGAGCGG </pre>
23406	<pre> ACCCCTGAGCTGCTTGGAGACTCGGCTGCCAGAGGGAGGGCATGGCCAGGCCCTGG CATTCACTGGCTCCACCCAGGCTGGCTGCCACAGTGCCCGGACACATTCA CCTCAGGCTCGTGTGGTGCAGGGACAAGAGGCTGTTGAGGCCAGGGACACCTGTGGAG GGCATTTCCCGTGGCCCCCGAGACCCCTAGATGGAGGAAGCGCTGCTGGGCCCTC </pre>

FIGURE 3J

TTACCGCTCACGGGAGCTGGGCCATGGATGGACAGGAGTCTTGTCCTGCTCAGCC
[C,T]
GGAGGCTGTGACGGCCCTGTCACAAGGTGACCTTGAGCACAGGCCGGGTGCCCC
AGGCTCGGCTCAGTTCTGGAGGTCAAGGGCATGGGTTGGGTAGTGGGTGGGAGGTGA
ATGTTTTCTAGAGATTCAAACGTGCTCCAGCAATTCTGTATAGTTTACCTCTGAGAAT
TACAATGTGAGAACCGCTCGATGTTGCATGTTCTGCGTACGTCCTGTGCTGCCCTGGCCG
TCAGGCCGGTGCTGCCCTGGTTGGCTGGACTTGGGCAGCCAGTGGGTGGCA

24078 GGCTGTGGAAGGCAGAGACGGTCCCTGGAGGCAGAGGAGCCAGGACAGCACCGTGACCCG
TGGAGCCGCCAGTGCCGGCAGTGTCTGGCCCTCCATAAGGGACGTATCCCTCTCAC
TGTGGCTGGGTTCTGTGGTAACTGTAACTACTGGTAAACGGCCTGTGTGCTT
CTCTCTGGTCTGGCTGGAGGGACGGGCTCAGGCCCTGAGCCCAGCGCTCCAGACAGGC
CTGTGCTGGTTCTGTAGGGAAATGGGTGTGGCGGGTCTGTGCCCTCCAGGACAGC
[G,A]
GCCATAGTGGACATGTGCTTAGACCTGTGTCCATGAGCCCCACTGCACCCCTGGCAAACA
GGGCCCTCCCGCTTGGCTGGCTGGAGATGGAGATGACAACGCCAAAGAACATTIG
GGGAAGAACCGCCATGCCAACGAGCAGAGTCAGAAGTCAGGAGGGATAGAATGAGCTTC
CCGTCCCCCCACCCACCCCTGTCTCTGAATCATGGCAAAACTAGCCTCCAGCCTCAG
CAGCTCACATGGGGACCGCACCCAAATCACCACAGGAAGGGTGGCCGGTCTCTGC

24777 GGAACATGGCAGGGTGGTCACCTTGTCTTCTCGAGACACTGGTGAGGTGTGGGTGCT
GTCTGGTCCCTTGATGCGCCCCCACACTGGGCAGAGTGGAGATGCTGTTGGGGGA
CATCAGCTCCACATCTGGCCAGAGGGAGCCCCGGGAAGGAAATGCTGAGGGGCCAGGG
CCTCGCCTGGATCTGCACAGCTTATAAGCAGCCAGGTGAGAGATGGGCTGTCTG
TGGTCCCAGAGACCCACGGCAGGAAATTCTGTCAACATCGGTGATGGCAGGGCCA
[G,A]
AGAACCGGTGGCACAAGGTGTCCTGGCTCTGCTCAACAAACAGCGAGTGCCCAGTGA
CTGGCACCAGGCCCTGCTTGGATGAGGACAACCGTCTGGAACGTCCACGCC
TATGAGACACAGCACGCCAGCACCGCAGTCACACACGGGGCTCGGTCAACCT
AGCTGCCCGGCTTGAGTGTGGCCTGGCTCTGTGAGCAGCGCCACCTGGGTGGCGGT
GGTGTGTGCTTCACTTCCACACACTCCCGTGCATGCTCCCCGGCCTTGTGGGTGGCTCG

Chromosome mapping:
Chromosome 10

FIGURE 3K